PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (http://bmjopen.bmj.com/site/about/resources/checklist.pdf) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Unexpectedly long hospital stays as an indicator of risk of unsafe
	care: an exploratory study.
AUTHORS	Borghans, Ine; Hekkert, Karin; den Ouden, Lya; Cihangir, Sezgin;
	Vesseur, Jan; Kool, Rudolf; Westert, Gert

VERSION 1 - REVIEW

REVIEWER	Ronald Lagoe Hospital Executive Council
	USA
REVIEW RETURNED	27-Feb-2014

GENERAL COMMENTS	This is not a major concern about methods, but it may deserve some additional explanation in the text. From the perspective of this reviewer, the development of the expected length of stay from which the "unexpectedly long length of stay" is calculated could use more explanation. In the United States, expected length of stays are computed based on comparison with benchmark populations. The authors might provide some additional material describing how these expected stays are identified in the Netherlands.
	This manuscript concerns and extremely important topic in health care, hospital patients with long lengths of stay. In the United States, much of the interest concerning this topic has been based on the need for increased efficiency of care. The authors of this manuscript have addressed the subject on the basis of patient outcomes. This is an important perspective that is consistent with the transition from volume to values in health care.
	The authors have developed their emphasis on long hospital stays as an outcomes indicator through discussions of adverse outcomes. The material in the First Paragraph on Page 5 of the manuscript is an excellent example of this.
	The data concerning the distribution of lengths of stay on Page 6 of the manuscript are interesting and could form the basis for comparisons with the experiences of other countries in subsequent studies.
	The request of this reviewer for additional explanation of "unexpected lengths of stay" is intended to support the structure of the study and should not be an obstacle to publication of the manuscript.

REVIEWER	Paul Aylin
	Imperial College London
REVIEW RETURNED	20-Mar-2014

This is a well written and clear paper outlining a proposed new patient safety measure, based on the percentage of patients with a prolonged length of stay of more than 50% of the expected length of stay. The methods are well described, and the authors acknowledge that the choice of 50% longer than the expected length of stay is just as arbitrary as choosing the 75th percentile. The findings are interesting, and the next step might be a comparison of complication codes in patients with UL-LOS, compared to patients under the threshold. Additionally, a case note review of those patients might also lend credibility to this measure as a proxy for complications. I have no changes to suggest.

REVIEWER	Oliver Groene London School of Hygiene and Tropical Medicine, UK
REVIEW RETURNED	24-Mar-2014

GENERAL COMMENTS

Abstract/Methods: the number of admissions/discharges that the calculation is based on should be mentioned. Abstract/Results: give details for the results of the correlation analysis

Background:

- Page 4, line 19: In my view figure 1 adds very little here and should be removed
- Page 4, line 38: "it is important to measure all three indicators...". Yes, I agree, but readmissions are not assessed in this paper

Methods:

- The various components of the methods section (page 5, line 36, page 7 line 19) are introduced as "first, ..", "second.." etc. This is confusing and should be removed, the subheadings are more than sufficient.
- If figure 2 is to be maintained, it should be designed more carefully.
- The definition of the threshold is not based on a strong justification, especially argument 2 (page 6, line 26) is a non-argument. I do agree that a certain amount of pragmatism is needed in derived thresholds and the table 1 provides more insight. However, it is unusual to present a Table in the methods section. This should be moved to either results or an annex.
- Is there information on the case-ascertainment rates among hospitals participating in the LMR? Page 7, line 50f "Five hospitals were excluded because they participated for less than 50%": does that mean that hospitals with 51-100% participation were included? Even a participation of 85%, while high, might induce substantial bias.
- Page 8, line 43: "All these additional criteria results in 58 hospitals remaining..". This is very important, but the process remains somewhat unclear. It should be supported by an appropriate flowchart showing all in and exclusion criteria and their effect on hospital participation.
- The indicator is calculated among survivors. The authors should comment on whether that introduces a bias whereby a patient suffering from a complication which leads to prolonged length of stay and eventually death would be excluded from the analysis. Could

that be better examined using a competing-risk model?

Results

- This section is very weak. From the results presented it is difficult to judge whether UL-LOS is a good indicator

Discussion

- Page 12, line 49: correlation with HSRM should not be confused with validity of UL-LOS
- Page 12, line 55: Yes, indeed, information on the sensitivity, specific and positive predictive value would be required to judge the validity of the indicator.

In conclusion, this paper represents in my opinion some exciting exploratory work that I hope will lead to further validation studies. However, at current, there are too many limitations with the study and I would recommend further work before publishing this manuscript. This might be important also in order to avoid creating expectations for the validity of a measure which, essentially, has not been established yet.

Unfortunately, this paper is somewhat disappointing. The validation is insufficient, writing could be much improved and the main expectation (answering the question of whether the UL-LOS is a good indicator) can not be answered with the current design. It would have been more appropriate had the authors presented this more as an exploratory study to highlight to need for careful validation of indicators based on routine data.

VERSION 1 – AUTHOR RESPONSE

Comment reviewer: 1

2) This is not a major concern about methods, but it may deserve some additional explanation in the text. From the perspective of this reviewer, the development of the expected length of stay from which the "unexpectedly long length of stay" is calculated could use more explanation. In the United States, expected length of stays are computed based on comparison with benchmark populations. The authors might provide some additional material describing how these expected stays are identified in the Netherlands.

In the methods section we changed the order of the first two items ('data' and 'definition') in order to make clear how we used the expected length of stay to generate the UL-LOS. In these sections we added additional (red colored) text to explain more clearly the expected length of stay. We hope we made clear to the reviewer that in the Netherlands the expected LOS also is computed based on comparison with benchmark populations.

We added reference 22 that shows recent data about extended lengths of stay of patients with complications.

Comment reviewer: 2

I have no changes to suggest.

Comments reviewer: 3

Title:

3) it is a bit long and should be shortened. I sympathize with the authors though regarding the need for careful wording

We removed the subtitle and on request of the editor (see comment 1) we added the study design. The whole title consists now of 15 words instead of 19 words.

Abstract/Methods:

4) the number of admissions/discharges that the calculation is based on should be mentioned. We added in the abstract/methods the number of discharges by adding the next sentence: "We used data of 61 Dutch hospitals. In total these hospitals had 1,400,000 discharges in 2011." In order to stay within the guidelines of maximum 300 words we removed the last sentence of the text under 'background'.

Abstract/Results:

5) give details for the results of the correlation analysis

We added the Pearson correlations between the UL-LOS indicator and the HSMR with all diagnoses groups and between the UL-LOS and discharging palliative patients with all diagnosis groups.

Background:

- 6) Page 4, line 19: In my view figure 1 adds very little here and should be removed We removed figure 1 and renumbered the next figures
- 7) Page 4, line 38: "it is important to measure all three indicators...". Yes, I agree, but readmissions are not assessed in this paper

At the moment the calculation of this indicator is being developed and therefore this outcome measure is still unknown in the Netherlands. Therefore we were not able to assess readmissions in this paper. We added in the discussion that it should be subject for future research.

Methods:

8) The various components of the methods section (page 5, line 36, page 7 line 19) are introduced as "first, ..", "second.." etc. This is confusing and should be removed, the subheadings are more than sufficient.

We removed the words 'first, etc'.

- 9) If figure 2 is to be maintained, it should be designed more carefully.
- We do agree with the reviewer that this figure needed layout editing. The word 'complications' was by mistake in Dutch and not all words fitted well in the arrows. We improved these issues.
- 10) The definition of the threshold is not based on a strong justification, especially argument 2 (page 6, line 26) is a non-argument. I do agree that a certain amount of pragmatism is needed in deriving thresholds and table 1 provides more insight. However, it is unusual to present a Table in the methods section. This should be moved to either results or an annex.

We do agree that argument 2 was not an argument for choosing a threshold of 50%. We moved this sentence some lines earlier in the text, moved table 1 to an annex, and renumbered the second table.

- 11) Is there information on the case-ascertainment rates among hospitals participating in the LMR? Page 7, line 50f "Five hospitals were excluded because they participated for less than 50%": does that mean that hospitals with 51-100% participation were included? Even a participation of 85%, while high, might induce substantial bias.

 See comment 12)
- 12) Page 8, line 43: "All these additional criteria results in 58 hospitals remaining..". This is very important, but the process remains somewhat unclear. It should be supported by an appropriate flowchart showing all in and exclusion criteria and their effect on hospital participation. We combined the comments 11 and 12 and added a flowchart in which all in- and exclusions are visualized and in which we also give insight in the group of hospitals that participated 51-100% of the year to the LMR.

13) The indicator is calculated among survivors. The authors should comment on whether that introduces a bias whereby a patient suffering from a complication which leads to prolonged length of stay and eventually death would be excluded from the analysis. Could that be better examined using a competing-risk model?

Under methods/definition we added the reason why we excluded patients who died in hospital. We excluded these patients because of the interrelationship between length of stay and mortality which makes it difficult to interpret length of stay data if non-survivors are included. We also added a reference to support our argument (PJ Marang-van de Mheen, 2014).

Discussion

- 14) Page 12, line 49: correlation with HSRM should not be confused with validity of UL-LOS We do agree with this, and therefore we don't make a link to the validity of the UL-LOS in this manuscript.
- 15) Page 12, line 55: Yes, indeed, information on the sensitivity, specificity and positive predictive value would be required to judge the validity of the indicator.

We do agree that further research is needed to judge the validity of the UL-LOS indicator. Therefore we added the next sentence in the discussion: "More research is needed to determine the validity of the UL-LOS indicator."

16) It would have been more appropriate had the authors presented this more as an exploratory study to highlight to need for careful validation of indicators based on routine data.

We agree with the reviewer and changed the subtitle to present it more as an exploratory study.

VERSION 2 – REVIEW

REVIEWER	Ronald Lagoe Reviewer
	Hospital Executive Council, USA
REVIEW RETURNED	24-Apr-2014

GENERAL COMMENTS	The comments of the authors concerning the issue that I (Reviewer 1) raised have been addressed adequately. These comments describe the authors' definition of an approach to expected length of stay, which is an important basis of the study. We also appreciate the additional reference addressing this subject that the authors have included.
	We have also reviewed the comments of Reviewer 3 and the responses of the authors to these comments. We believe that these responses help clarify the study without detracting from the principal objective and related information.
	As indicated in earlier comments, we believe that the subject of this manuscript, the long lengths of stay as an indicator of patient outcomes is extremely important and worthy of publication.

REVIEWER	Oliver Groene
	London School of Hygiene and Tropical Medicine, UK
REVIEW RETURNED	07-May-2014

GENERAL COMMENTS	I am delighted that you found my comments useful. As a result, I believe the revised manuscript has improved and now clearly communicates on your interesting developmental work on a new
	indicator of unsafe care. The results are of interest internationally and will hopefully lead to further validation studies in the future.